

Input file F1h14273new; Duput File F1h14273tra
Sequence length 1743

TCGCGACTAGTTCTAGCCGCTGGGGCCGCCAGCGCCGGGA	ATG	TCC	CCT	GAA	TGC	E	C	A	R	A	A	GCG	9
G D A P L R S L E Q A N R T R F P F S												27	
GGC GAC GCG CCC TTG CGC AGC CTG GAG CAA GCC AAC CGC ACC CGC TTT CCC TTC TTC TCC												29	
D V K G D H R L V L A A V E T T V L V L												49	
GAC GTC AAG GGC GAC CAC CGG CTG GTG CTG GCG GTG GAG ACA ACC GTG CTG GTG CTC												147	
I F A V S L L G N V C A L V L V A R R R												69	
ATC TTT GCA GTG TCG CTG CTG GGC AAC GTG TGC GGC CTG GTG CTG GCG CGC CGA CGA												207	
R R G A T A C L V L N L F C A D L L F I												89	
CGC CGC GGC GCG ACT GCC TGC CTG GTA CTC AAC CTC TTC TGC GCG GAC CTG CTC TTC ATC												267	
S A I P L V L A V R W T E A W L L G P V												109	
AGC GCT ATC CCT CTG GTG CTG GGC GTG GCG TGG ACT GAG GCC TGG CTG CTG GGC CCC GTT												327	
A C H L L F Y V M T L S G S V T I L T L												129	
GCC TGC CAC CTG CTC TTC TAC GTG ATG ACC CTG AGC GGC AGC GTC ACC ATC CTC ACG CTG												387	
A A V S L E R M V C I V H L Q R G V R G												149	
GCC GCG GTC AGC CTG GAG CCG ATG GTG TGC ATC GRG CAC CTG CAG CGC GGC GTG CCG GGT												447	
P G R R A R A V L L A L I W G Y S A V A												169	
CCT CGG CGG GCG GCG GCA GTG CTG CTG GCG CTC ATC TGG GCC TAT TCG GCG GTC GCC												507	
A L P L C V F F R V V P Q R L P G A D Q												189	
GCT CTG CCT CTC TGC GTC TTC TTT CGA GTC CTC CCG CAA CGG CTC CCC GGC GGC GAC CAG												567	

TO FIG. 1B.

FIG. 1A.



1/15

400755987.072202

COPY OF PAPERS
ORIGINALLY FILED



10075687.072202

2/15

COPY OF PAPERS
ORIGINALLY FILED

FROM FIG. 1A.

E	I	S	I	C	T	L	I	W	P	T	I	P	G	E	I	S	W	D	V	209
GAA	ATT	TCG	ATT	TGC	ACA	CTG	ATT	TGG	CCC	ACC	ATT	CCT	GGA	GAG	ATC	TCG	TGG	GAT	GTC	627
S	F	V	T	L	N	E	L	V	P	G	L	V	I	V	I	S	Y	S	K	229
TCT	TTT	GTT	ACT	TTG	AAC	TTC	TTG	GTG	CCA	GGA	CTG	ATT	GTG	ATC	AGT	TAC	TCC	AAA	687	
I	L	Q	I	T	K	A	S	R	K	R	L	T	V	S	L	A	Y	S	E	249
ATT	TTA	CAG	ATC	ACA	AAG	GCA	TCA	AGG	AAG	AGG	CTC	ACC	GTA	AGC	CTG	GCC	TAC	TCG	GAG	747
S	H	Q	I	R	V	S	Q	Q	D	F	R	L	F	R	T	L	F	L	L	269
AGC	CAC	CAG	ATC	CGC	GTG	TCC	CAG	CAG	GAC	TTC	CGG	CTC	TTC	CGC	ACC	CTC	TTC	CTC	CTC	807
M	V	S	F	F	I	M	W	S	P	I	I	I	T	I	L	L	I	L	I	208
ATG	GTC	TCC	TTC	ATC	ATG	ATG	TGG	AGC	CCC	ATC	ATC	ATC	ACC	ATC	CTC	CTC	ATC	CTG	ATC	867
Q	N	F	K	Q	D	L	V	I	W	P	S	L	F	F	W	V	V	A	F	309
CAG	AAC	TTC	AAG	CAA	GAC	CTG	GTC	ATC	TGG	CCG	TCC	CTC	TTC	TTC	TGG	GTG	GTG	GCC	TTC	927

TO FIG. 1C.

FIG. 1B.



10075987.072202

3/15

COPY OF PAPERS
ORIGINALLY FILED

FROM FIG. 1B.

T F A N S A L N P I L Y N M T L C R N E 329
ACA TTT GCT AAT TCA GCC CTA AAC CCC ATC CTC TAC AAC ATG ACA CTG TGC AGG AAT GAG 987
W K K I F C C F W F P E K G A I L T D T 349
TGG AAG AAA ATT TTT TGC TGC TTC TGG TTC CCA GAA AAG GGA GCC ATT TTA ACA GAC ACA 1047

S V K R N D L S I I S G . 362
TCT GTC AAA AGA AAT GAC TTG TCG ATT ATT TCT GGC TAA 1086

TTTTCTTTATAGCCGAGTTTCTCACACCTGGCGAGCTGTGGCATGCTTTTAAACAGAGTTTCATTTCCAGTACCCCTCCA
TCAGTGCACCCCTGCTTTAAGAAAAAGAACCTATGCAATAGACATCCACAGCGTCGGTAAATTAAGGGGTGATCACCAA
GTTTCATAATATTTTCCCTTTTATAAAGGATTTTGTGGCCAGGTGCAGTGGTTTCATGCTGTATCCAGCAGTTTGGG
AGGCTGAGGTGGTGGATCACCTGAGGTGAGGAGTTCGAGACCAACCTGACCAACATGGTGAGACCCCGTCTCTACTA
AAAAATAAAAAAAAATTAGCTGGGAGTGGTGGGACCTGTAAATCCTAGCTACTTTGGGAGGCTGAACCCAGGAGAA
CTCTTGAACCTGGGAGGCAGAGTTGCAGTGAGCCGAGATCGTGCCATTGCACCTCAACAGGGCAACAAGAGTGAAAC
TCCATCTTAAAAAAAATAAAGATTTGTTATGGGTTCCTTTTAAATGTGAACTTTTTTAGTGTTGTGTAATATG
ATCAAAATTAATAAATATTTATTTATGACTGTTTCAGCAAAAAAAAAGGGGGG

RTA01/2057957v1

FIG. 1C.

10075987 .072202

**COPY OF PAPERS
ORIGINALLY FILED**



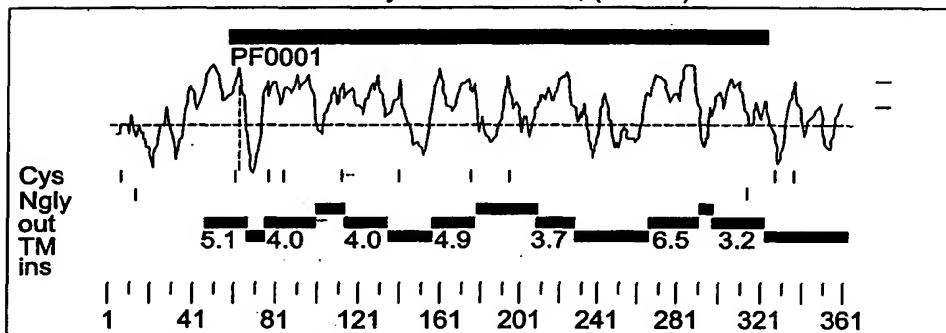


10075987.072202

6/15

COPY OF PAPERS
ORIGINALLY FILED

Analysis of Flh14273, (362 aa)



>Flh14273, 1086 bases, 1825 checksum.
MSPECARAAQDAPLRSLSEQANRTRFPFFSDVKGDHRLVLAAVETTVLVLEFAVSLLGNVC
ALVLVARRRRRGATAQLVLNLFCA DLLFYSAIPLVLAVRVTEAVLLGPVACHLLFYVMTL
SGSVTILTLAAVSLERMVCI VHLQRCVRGPGRRARAVLLALIWGYSAAALPLCVFFRVV
PQRLPGADQEISICTLIWPTIPGEISVDVSFVTLNFLVPLVIVIVISYSKILQITKASRR
LTVSLAYSESHQIRVSQDDFLFRTLFLLMVSFFIMWSPITITILLILIQNFKQDLVIWP
SLFFVVAFTFANSALNPILYNMTLCRNEWKKIFCCPWFGEKGAITDTSVKRNDLSIIS
G+

FIG. 4.



10075987.072202

7/15

COPY OF PAPERS
ORIGINALLY FILED

Prositate Pattern Matches for F1h14273

>PS00001/PD0C00001/ASN_GLYCOSYLATION N-glycosylation site.

Query: 21 NRTR 24

Query: 322 NMTL 325

>PS00004/PD0C00004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site

Query: 239 KRLT 242

>PS00005/PD0C00005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 237 SRK 239

Query: 350 SVK 352

>PS00006/PD0C00006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 256 SQQD 259

>PS00008/PD0C00008/MYRISTYL N-myristoylation site.

Query: 57 GNVCAL 62

Query: 72 GATACL 77

Query: 343 GAILTD 348

>PS00009/PD0C00009/AMIDATION Amidation site.

Query: 150 PCRR 153

>PS00029/PD0C00029/LEUCINE_ZIPPER Leucine zipper pattern.

Query: 106 LGPVACHLLFYVMTLSGVSITL 127

FIG. 5.



8/15

COPY OF PAPERS
ORIGINALLY FILED

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
46	66	out->ins	5.1
75	98	ins->out	4.0
113	134	out->ins	4.0
156	177	ins->out	4.9
209	227	out->ins	3.7
266	289	ins->out	6.5
297	321	out->ins	3.2

>F1h14273,

MSPECARAAGDAPLRSLSEQANRTRFPFSDVKGDHRLVLAAVETTVLVLIPAVSLLGNVC
ALVLVARRRRRGATACLVLNLF CADLLFISAWPLVLAVRWTEAWLLGPVACHLLFYVMTL
SGSVTILTLAAVSLERMVCIVHLQRGV RGPGRRARAVLLALIWGYSAAALPLCVFFRVV
PQRLPGADQEISICTLIWPTIPGEISWDVSFVTLNFLVPLVIVISYSKILQITKASRKRL
LTVSLAYSESHQIRVSQQDFRLFRTLFLLMVSFFIMWSPITITILLILIQNFKQDLVIWP
SLFFWVVAFTFANSALNPILYNMTLCRNEWKKIFCCWFPEKGAILTDSVKRNDLSIIS
G

Transmembrane Segments for Presumed Mature Peptide

Start	End	Orient	Score
14	37	ins->out	4.0
52	73	out->ins	4.0
95	116	ins->out	4.9
148	166	out->ins	3.7
205	228	ins->out	6.5
236	260	out->ins	3.2

>F1h14273, _mature

LVLVARRRRRGATACLVLNLF CADLLFISAIPLVLAVRWTEAWLLGPVACHLLPYVMTLS
GSVTILTLAAVSLERMVCIVHLQRGV RGPGRRARAVLLALIWGYSAAALPLCVFFRVVP
QRLPGADQEISICTLIWPTIPGEISWDVSFVTLNFLVPLVIVISYSKILQITKASRKRL
TVSLAYSESHQIRVSQQDFRLFRTLFLLMVSFFIMWSPITITILLILIQNFKQDLVIWPS
LFFWVVAFTFANSALNPILYNMTLCRNEWKKIPCCWFPEKGAILTDSVKRNDLSIISG

FIG. 6.



10075987.072202

COPY OF PAPERS
ORIGINALLY FILED

9/15

Input file 14273mj Output File 14273mtra
Sequence length 1560

TTGCCAAGCTCAGCGTAAGCCTCTTCCACTGCAATCTCACAGAAGGGTTTCATGGAGTGCTTCACACCATCAGTGACCA
CTCCAGACTTGTCCGGCTTTACCCGAATCTTCACAGCGGAGTCGATGACCCCTCTTGACAGCCACGAGCGCGCGCAGCTC
CGCCATCTTCCCGGACGCGTGGCGCGGCGCGGC ATG TCC CCT GAG TGT GCA CAG ACG ACG GGC 10
P G P S H T L D Q V N R T H F P F F S D 30
CCT GCT CCC TCG CAC ACC CTG GAC GAA GTC AAT CGC ACC CAC TTC CCT TTC TTC TCG GAT 30
V K G D H R L V L S V V E T T V L G L I 50
GTC AAG GGC GAC CAC CGG TTG GTG AGC GTC GTG GAG ACC ACC GTT CTG GGA CTC ATC 150
F V V S L L G N W C A L V L V A R R R 70
TTT GTC GTC TCA CTG CTG AAC TTG TGT GCT CTA GTG CTG GTG GCG CGC CGT CGG CGC 210
R G A S A S L V L N L F C A D L L F T S 90
CGT GGG GCG TCA GCC AGC CTG GTG CTC AAC CTC TTC TGC GCG GAT TTG CTC TTC ACC AGC 270
A I P L V L V V R W T E A W L L G P V 110
GCC ATC CCT CTA GTG CTC GTG GCG TGG ACT GAG GCC TGG CTG TTG GGG CCC GTC GTC 330
C H L L F Y V M T M S G S V T I L T L A 130
TGC CAC CTG CTC TAC GTG ATG ACA ATG AGC AGC AGC GTC ACG ATC CTC ACA CTG GCC 390
A V S L E R M V C I V R L R R G L S G P 150
GCG GTC AGC CTG GAG CGC ATG GTG TGC ATC GTG CGC CTC CGG CGC GGC TTG AGC GGC CCG 450
G R R T Q A A L L A F I W G Y S A L A A 170
GGG CGG CGG ACT CAG GCG GCA CTG CTG GCT TTC ATA TGG GGT TAC TCG GCG CTC GCC GCG 510

TO FIG. 7B.

FIG. 7A.



10/15

FROM FIG. 7A.

L P L Y I L F R V V V P Q R L P G G D Q E
 CTG CCC CTC TAC ATC TTG TTC CGC CGC CCG CAG CGC CTT CCC GGC GGC GAC CAG GAA
 I P I C T L L D W P N R I G E I S W D V F
 ATT CCG ATT TGC ACA TTG GAT TGG CCC AAC CGC ATA GGA GAA ATC TCA TGG GAT GTG TTT
 F E T L N F L V P G L V I V I S Y S K I
 TTT GAG ACT TTG AAC TTC CTG CTG CCG GGA CTG GTC ATT GTG ATC AGT TAC TCC AAA ATT
 L Q I T K A S R K R L T L S L A Y S E S
 TTA CAG ATC ACG AAA GCA TCG CCG AAG AGG CTT ACG CTG AGC TTG GCA TAC TCT GAG AGC
 H Q I R V S Q Q D Y R L F R T L F L L M
 CAC CAG ATC CGA GTG TCC CAA CAA CAC TAC CGA CTC TTC CGC AGC CTC TTC CTG CTC ATC
 V S F F I M W S P I I I T I L L I L I Q
 GTT TCC TTC TTC ATC AIG TGG AGT TCC ATC ATC ACC ATC CTC CTC ATC TTG ATC CAA
 N F R Q D V I V P S L F F W V A F T
 AAC TTC CGG CAG GAC CTC ATC TGG CCA TCC CTT TTC TGG GTG GTG GCC TTC ACG
 F A N S A L N P I L Y N M S L F R N E W
 TTT GCC AAC TCT GCC CTA AAC CCC ATA CTG TAC AAC ATG TCG CTG TTC AGG AAC GAA TGG
 R K I F C C F F P E K G A I F T D T S
 AGG AAG ATT TTT TGC TGC TTC TTT TTT CCA GAG AAG GGA GCC ATT TTT ACA GAT ACG TCT
 V R R N D L S V I S S
 GTC AGG CGA AAT GAC TTG TCT GTT ATT TCC AGC TAA

362
1086

CTAGCCTCTGGTGCCAGGTGAACCACGGGTGTGCATGTAAAGCGAGTTAACTTCAAGGAAAGCCACCAGTGGCCCTGC
 TTTAAAAATACCCGACTTCCACACAGCAGGCATCTACGGAGCCAGCAAAATTAAGGAATGATCGCTCAGTATAAAAAATTT
 TTTCCCTAAAAGAACTTCTATGGGTTCCTTTTGTGAACCTTTTAAAGTGTGTTGTAATATGATCTAGTTAATAAATT
 TTTATTATAACGTGTTCTTACAAAAAATTTTTTTTTT

FIG. 7B.



11/15

COPY OF PAPERS
ORIGINALLY FILED

Query: 14273m,

Scores for sequence family classification (score includes all domains):

Sequence	Description	Score	E-value	N
7tm_1	PF00001 7 transmembrane receptor (rhodopsin)	118.8	1e-35	1

Parsed for domains:

Sequence	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value	-
7tm_1	1/1	57	321	..	1 259 []	118.8	1e-36	

Alignments of top-scoring domains:

7tm_1: domain 1 of 1, from 57 to 321: score 118.8, E = 1e-36

*->GNILVilvllrtkkllrtptnifllNLAvADLLflltppwalyyllvg

GN+ +++++r +tr ++ +NL ADLLf+ +p++ ++ ++

14273m, 57 GNVCALVLVAR-RRRRGASASLVNLF CADLLFSAIPLVLVVR-WT 101

gaadWpfGaalCk1vta1dvvnmyaSi111Lta1SiDRY1A1vhP1yrrr

e W++G+++C+1+ ++++++ + 11+L+a S++R + Iv 1r +

14273m, 102 --EAWLLGPVCHLLPYVMTMSGSVTILTLAAVSLERMVCIV-RLRRGLS 148

rtsprA.kvvi11vWv1a111s1Pp11fswvktveagngtlnvntvC1

rr+++++++W ++1++1P ++++ v + ++g ++ +C+

14273m, 149 GP-GRRTqAALLAFIWGYSALAALPLYILFRVVPQRLPGGD--QEITPCT 195

idfpeestasvstwlrsvv11st1vgF21P11v11voYtr11rtlr...

+d+p++ + ++s+ +++ ++ F1+P 1v1++ Y+ 11 + + ++

14273m, 196 LDWPNRIG-----EISWDVFFETLNFLVPLVIVISYSKILQITKasrk 239

...kaakt11vvvvvFv1CW1Pyfiv111dt-

+++++ +t1++++v F++ W P i++11 +

14273m, 240 r1t1slayseshqirvsqdyRLFRTLFLLMVSFFIMWSPIIITILLILI 289

c.1a11nestCelervlptallvt1wLayvNsc1NP11Y<-*

++ + + p++++ +++++N+1NP1+Y

14273m, 290 QnFR-----QDLVIVPSLFFVWVAFTFANSALNPILY 321

FIG. 8.



10075087.072202

COPY OF PAPERS
ORIGINALLY FILED

12/15

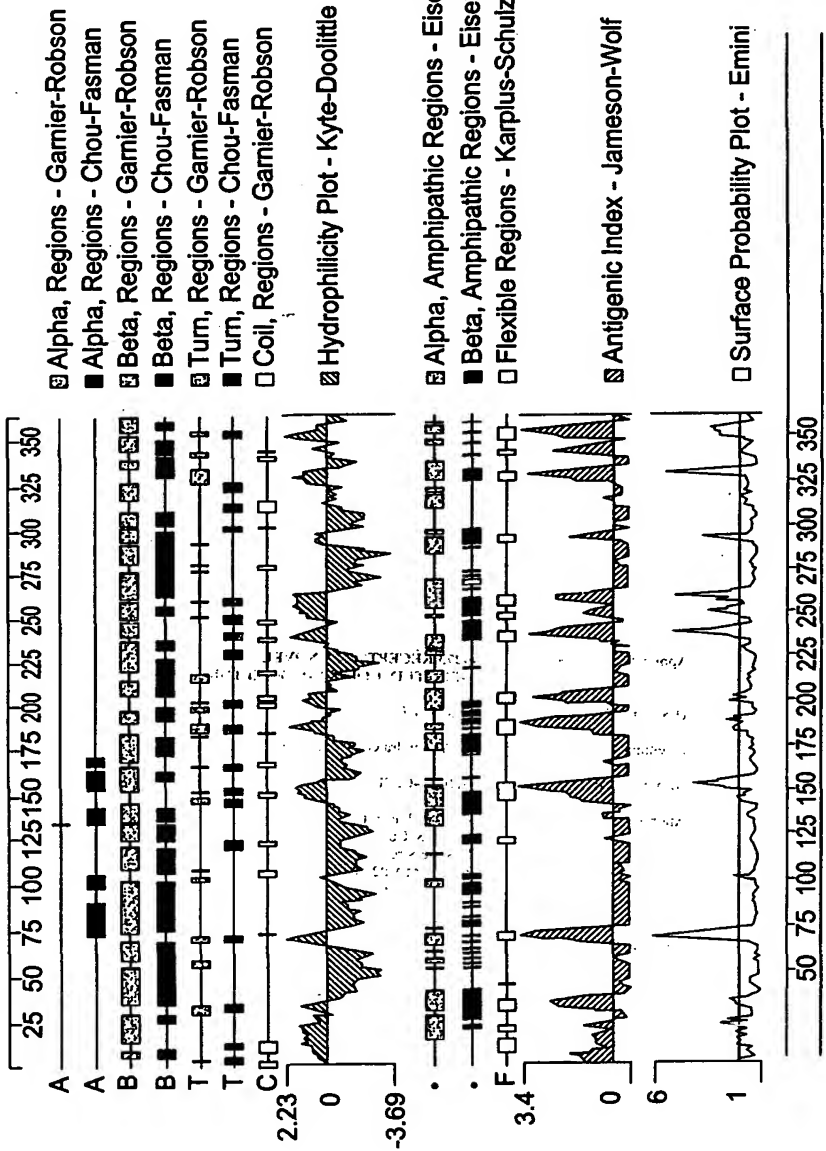


FIG. 9.

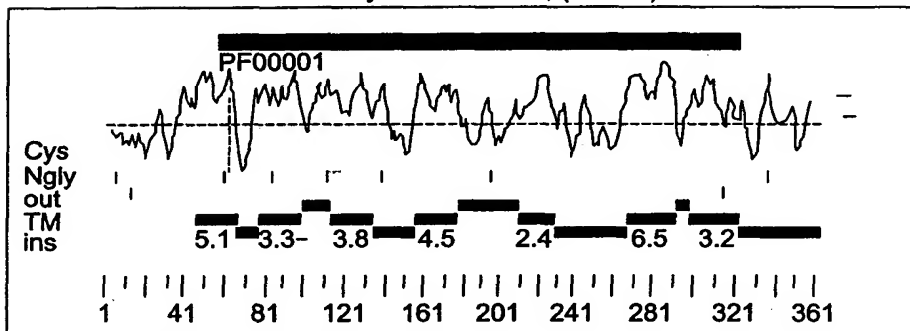


10075987.072202

13/15

COPY OF PAPERS
ORIGINALLY FILED

Analysis of 14273m, (362 aa)



>14273m, 1086 bases, 6943 checksum.
MSPECAQTGTGCPSTLDQVNRTHFPFFSDVKGDRHLVLSVVETTVLGLIFVVSLLGNVC
ALVLVARRRRRGASASLVNLFCADLLFTSAIPLVLVVRWTEAWLLGPVVCHLLFVMTM
SGSVTILTLAAVSLERMVCIVRLRRLSGPGRRTQAALLAFIWGYSALAALPLYILFRVV
PQRLPGGDQEIPICTLDWPNRIGEISWDVFFETLNLFLVPLVIVISYSKILQITKASRRR
LTLSLAYSESHQIRVSQQDYRLFRTLFLMVSTFVWSPITITILLITQNRQDLVIWP
SLFFVWVAFTFANSALNPILYNMSLFRNEWKIFCCFPPEKGAIFTDTSVRRNDLSVIS
S*

FIG. 10.



10075987.072202

14/15

COPY OF PAPERS
ORIGINALLY FILED

Prosite Pattern Matches for 14273m,

>PS00001/PDOC0001/ASN_GLYCOSYLATION N-glycosylation site.

Query: 21	NRTH	24
Query: 322	NRTH	325

>PS00002/PDOC0002/GLYCOSAMINOGLYCAN Glycosaminoglycan attachment site.

RU	Additional rules:
RU	There must be at least two acidic amino acids (Glu or Asp) from -2 to -4 relative to the serine.

Query: 148	SGPG	151
------------	------	-----

>PS00004/PDOC0004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation

Query: 239	KRLT	242
------------	------	-----

>PS00005/PDOC0005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 237	SRK	239
Query: 350	SVR	352

>PS00006/PDOC0006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 40	SVVE	43
Query: 256	SQQD	259

>PS00008/PDOC0008/MYRISTYL N-myristoylation site.

Query: 57	GNVCL	62
Query: 72	GASASL	77
Query: 343	GAIFTD	348

>PS00009/PDOC0009/AMIDATION Amidation site.

Query: 150	PGRR	153
------------	------	-----

FIG. 11.



Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
46	66	out->ins	5.1
77	98	ins->out	3.3
113	134	out->ins	3.8
156	177	ins->out	4.5
209	227	out->ins	2.4
266	289	ins->out	6.5
297	321	out->ins	3.2

>14273m,

MSPECAQTTGPGPSHTLDQVNRTHFPFFSDVKGDHRLVLSVVETTTLGLIFVVSLLGNVC
ALVLVARRRRRGASASLVNLFCADLLFTSAIPLVLVVRWTEAWLLGPVVCHLLFYVMTM
SGSVTILTLAAVSLERMVCIVLRRGLSGPGRRTQAALLAFIWGYSALAALPLYILFRVV
PQRLPGGDQEIPICTLDWPNRIGEISWDVFFETLNLVPLVIVISYSKILQITKASRR
LTLSESESHQIRVSQQDYRLFRTLFLLMVSFFIMWSPIIITILLILIQNFRQDLVIWP
SLFFWVVAFTFANSALNPILYMMSLFRNEWKIFCCFFPEKGAIFTDTSVRRNDLSVIS
S

Transmembrane Segments for Presumed Mature Peptide

Start	End	Orient	Score
16	37	ins->out	3.3
52	73	out->ins	3.8
95	116	ins->out	4.5
148	166	out->ins	2.4
205	228	ins->out	6.5
236	260	out->ins	3.2

>14273m, _mature

LVLVARRRRRGASASLVNLFCADLLFTSAIPLVLVVRWTEAWLLGPVVCHLLFYVMGMS
GSVTILTLAAVSLERMVCIVLRRGLSGPGRRTQAALLAFIWGYSALAALPLYILFRVVP
QRLPGGDQEIPICTLDWPNRIGEISWCVFFETLNLVPLVIVISYSKILQITKASRR
LTLSESESHQIRVSQQDYRLFRTLFLLMVSFFIMWSPIIITILLILIQNFRQDLVIWPS
LFFWVVAFTFANSALNPILYMMSLFRNEWKIFCCFFPEKGAIFTDTSVRRNDLSVIS

FIG. 12.